SENSITIVITY COMPUTATION OF PERIODIC AND CHAOTIC LIMIT CYCLE OSCILLATIONS

QIQI WANG * AND RUI HU *

Abstract. Sensitivity analysis can be a powerful tool for design optimization, error analysis and data assimilation in simulating engineering systems. For computing sensitivity of time averaged quantities in system exhibiting chaotic oscillations, conventional sensitivity analysis methods (tangent linear and adjoint methods) can produce erroneous results. This paper presents a new method for computing sensitivity derivatives for systems that exhibit both periodic and aperiodic (chaotic) limit cycle oscillations. This method overcomes the growth of sensitivity that disables traditional sensitivity analysis methods.

Key words. Sensitivity analysis, linear response, adjoint equation, unsteady adjoint, chaos, statistical average, Lyapunov exponent, Lyapunov covariant vector, Lorenz attractor.

AMS subject classifications.

1. Introduction.

Adjoint based computational sensitivity analysis of flow fields has become a powerful technology in designing airplanes [1, 2]. In an adjoint based analysis, sensitivity to small perturbations is traced back from the quantity of interest, such as lift or drag coefficient, to perturbations to the flow field. The adjoint equations govern the backward propagation of sensitivities, and are derived from the governing equations of the flow. The solution to the adjoint equations reveals the gradient of the quantity of interest with respect to perturbations to the flow field, e.g., from modifying the surface geometry. This gradient can then be used to drive a gradient based optimization scheme to maximize or minimize the quantity of interest. The adjoint method has been shown to be able to solve very high dimensional optimization problems typically found in aerodynamic shape design and optimal flow control applications [3, 4].

In unsteady aeroelastic problems, adjoint based methods can be a powerful and transformative technology. Sensitivity computed from such analysis can be used to design non-conventional airplanes, rotorcrafts and micro air vehicles to maximize performance and minimize cost and risk [5, 6, 7, 8]. In computing sensitivities of time-averaged quantities in unsteady aeroelastic simulations, many quantities of interest are time-averaged quantities, such as aerodynamic forces and structural fatigue. Despite recent advances in algorithms for sensitivity analysis unsteady fluid flows [9, 10, 11, 7, 12, 13, 14], sensitivity analysis of long time averages has shown to yield erroneous results for chaotic dynamical systems [15]. As illustrated in [15] and in Section 3 of this paper, naively computed sensitivity in a chaotic dynamical system grows as $O(e^{\lambda T})$ as the averaging length T increases, where λ is the maximal Lyapunov exponent and has a positive value.

The challenge of computing sensitivity of long time averages in chaotic dynamical systems has attracted much interest. The ensemble adjoint approach [15, 16] approximates the sensitivity gradient by combining many adjoint evaluations on randomly sampled trajectories. However, the required number of trajectories can be very large, making the method computational costly even for small dynamical systems like the Lorenz attractor. Method based on linearizing the Fokker-Planck equation [17] computes sensitivity of long time averages by solving a PDE in the phase space. However,

*Department of Aeronautics and Astronautics, MIT, 77 Mass Ave, Cambridge, MA 02139, USA

a small random noise perturbation must be added to the dynamical system, whose effect is difficult to quantify. Methods based on the Fluctuation-Dissipation Theorem has proven successful on some dynamical systems [18]. However, they require the equilibrium distribution of the dynamical system to have properties similar to that of the Gaussian distribution. The recently proposed method based on Lyapunov eigenvector decomposition [19] computes accurate sensitivity on a short trajectory. However, the computational cost is linear to the number of positive Lyapunov exponents, making it expensive when many positive Lyapunov exponents coexist.

This paper present a new computational method based on solving the least squares problem of the linearized equations. The solution of the least squares problem reveals a pair of shadowing trajectories. The algorithm is easy to implement. The computation cost of this new method is independent of the number of positive Lyapunov exponents, making it potentially suitable for high dimensional chaos.

The rest of this paper is organized as follows. Section 2 describes the conventional methods for sensitivity analysis, including the tangent linear and the adjoint methods. Section 3 describes the challenge of using traditional sensitivity analysis methods for chaotic oscillations. Section 4 outlines our new method for computing sensitivity of both periodic and chaotic oscillations. Sections 5, 6 and 7 apply this new method to the van der Pol oscillator, the Lorenz attractor and an aero-elastic oscillation model. Section 8 concludes the paper.

2. Conventional method for sensitivity computation. Conventional tangent linear method for sensitivity computation [20] can apply to a autonomous dynamical system [1]

$$\frac{du}{dt}(t,s) = f(u(t,s),s) \tag{2.1}$$

parameterized by s, which can represent control variables, design variables, or uncertain parameters. The initial condition of the dynamical system is given as

$$u(0,s) = u_0(s)\ . \tag{2.2}$$

Conventional tangent linear method [20] of sensitivity analysis computes the derivative of an objective function

$$\mathbf{J}(s) = \int_0^T J(u(t,s),s)\,dt \tag{2.3}$$

to the parameter s.

The conventional tangent linear method starts by solving an initial value problem of the tangent equation

$$\frac{dv_{iv}}{dt} = \frac{\partial f}{\partial u}v_{iv} + \frac{\partial f}{\partial s}\ , \quad v_{iv}(0) = \frac{du_0}{ds} \tag{2.4}$$

where v_{iv} represents the linearized state

$$v_{iv}(t,s) = \frac{\partial u}{\partial s}(t,s) \tag{2.5}$$

[1] The conventional tangent and adjoint methods work for non-autonomous systems. However, we focus on autonomous systems in this paper.

of the initial value problem. The desired sensitivity derivative can then be computed via chain rule:

$$\frac{d\mathbf{J}}{ds} = \int_0^T \left(\frac{\partial J}{\partial u} v_{iv} + \frac{dJ}{ds} \right) dt \tag{2.6}$$

This method is also known as the forward sensitivity analysis method. Another popular method for sensitivity computation is the adjoint method [20]. It is the formal dual of the tangent linear method. The adjoint method has the advantage of being able to simultaneously compute the derivative of an objective function to many parameters. For this reason, it is often used in optimal control, inverse problems and data assimilation.

3. Breakdown of the conventional method for time averaged output quantities. For a dynamical system exhibiting periodic and chaotic limit cycle oscillations, many output variables are long time averaged output quantities, which represent long term behavior of the dynamical system.

$$\langle J \rangle(s) = \lim_{T \to \infty} \frac{1}{T} \int_0^T J(u(t,s),s)\, dt\ , \tag{3.1}$$

Under the ergodic assumption, such infinite time averages are independent of the initial condition $u(0,s)$. They only depend on the parameter s. These quantities are often called the "climate" of the dynamical system. Many output quantities of key interest in aeroelastic oscillations, such as mean magnitude of oscillation and structural fatigue, can be approximated by long time averages.

In a computational simulation, an infinite time average must be approximated with a finite time average

$$\langle J \rangle(s) \approx \overline{J}(s) = \frac{1}{T} \int_0^T J(u(t,s),s)\, dt\ . \tag{3.2}$$

Because $\overline{J}(s)$ is in the form of Equation (2.3), the conventional method described in Section 2 can be used to compute $d\overline{J}(s)/ds$. One might think that since $\overline{J}$ is a good approximation to $\langle J \rangle$ when T is large, the computed sensitivity derivative $d\overline{J}/ds$ should be a good approximation to the desired sensitivity derivative $d\langle J \rangle/ds$ when T is large enough. However, this is not the case.

The conventional sensitivity analysis method described in Section 2 breaks down for long time averaged output quantities because $d\overline{J}(s)/ds$ can be a terrible approximation to $d\langle J \rangle(s)/ds$. They can be orders of magnitude different for large T. The finite time average approximation contains a small error $\overline{J}(s) - \langle J \rangle(s)$. The magnitude of this error is small for large T; however, its derivative to s can increase as T increases. As a result, the derivative $d\overline{J}(s)/ds$ does is not a useful approximation of $d\langle J \rangle(s)/ds$, rather it grows exponentially with increasing T [15]. If one decompose $J(u(t,s),s)$ into its time average and a zero-mean fluctuating term

$$J(u(t,s),s) = \langle J \rangle(s) + \widetilde{J}(u(t,s),s),$$

the sensitivity derivative of $\overline{J}(s)$ to s can be written as

$$\frac{d\overline{J}(s)}{ds} = \frac{d\langle J \rangle(s)}{ds} + \frac{1}{T} \int_0^T \frac{d\widetilde{J}(u(t,s),s)}{ds}\, dt\ ,$$

while the first term is the desired sensitivity derivative, the second term often does not decay to 0 as T increases [15] [11]. In particular, for chaotic dynamical systems, it is observed that $\partial u(t,s)/\partial s$ grows exponentially as t increases, due to the "butterfly effect"[21], also known as "chaotic error growth"[15]. As a result,

$$\frac{d\overline{J}(s)}{ds} - \frac{d\langle J\rangle(s)}{ds} = \frac{1}{T}\int_0^T \frac{d\widetilde{J}(u(t,s),s)}{ds}\,dt = \frac{1}{T}\int_0^T \frac{\partial \widetilde{J}}{\partial s} + \frac{\partial \widetilde{J}}{\partial u}\frac{\partial u(t,s)}{\partial s}\,dt$$

increases exponentially as T increases. The $d\overline{J}(s)/ds$ computed with either the conventional tangent linear method or the adjoint method diverges from the desired sensitivity $d\langle J\rangle(s)/ds$.

This divergence behavior is examined by Lea et al. [15] in studying the Lorenz system

$$\frac{dx}{dt} = \sigma(y-x)\ , \quad \frac{dy}{dt} = x(\rho - z) - y\ , \quad \frac{dz}{dt} = xy - \beta z\ . \tag{3.3}$$

They observe that when T increases, the difference between $d\overline{z}/d\rho$ and $d\overline{z}/d\rho$ does not decrease but increases exponentially fast. Following Lea et al.'s numerical experiment, we compute $\overline{z}$ by time averaging z computed from the initial condition $x = 1, y = 1, z = 30$ over a time period of $T = 2.26$ and 131.4, respectively. The sensitivity derivative of $\overline{z}$ is computed with the procedure in Section 2, where the discrete linearization is used. Figure 3.1 shows the sensitivity derivative of the finite time averages for different averaging lengths T. The divergence of sensitivities is a fundamental challenge in computing sensitivities for time averaged output quantities in unsteady oscillations.

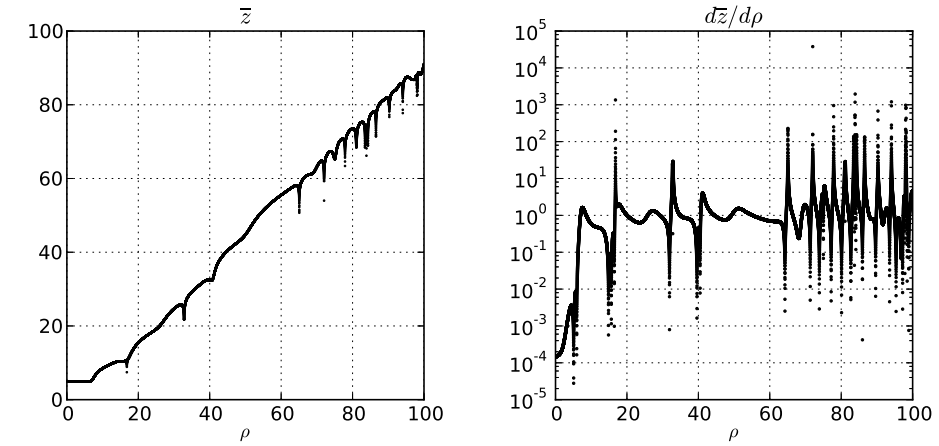
4. Least Squares Sensitivity Analysis of time averaged output quantities. Our new method takes advantage of the ergodic assumption of the dynamical system. The linearized equation (2.4) approximates the difference between two adjacent trajectories, one with a perturbation in the right hand side, and one without. The time averaged objective function we try to capture does not depend on the initial condition. Therefore, the initial condition of the linearized equation, describing the initial difference between the two adjacent trajectories, can take an arbitrary value. Our new method solves the linearized equation (2.4) without an explicit initial condition. The freed-up degrees of freedom resulting from the lack of initial condition are used to explicitly control the growth of the solution of the linearized equation, thus suppressing the butterfly effect.

4.1. Computing sensitivity of time averaged output from tangent direction of a shadowing trajectory. Consider a semi-infinite trajectory $u(t), t \in [0,\infty)$ satisfying

$$\frac{du}{dt} = f(u,s) \tag{4.1}$$

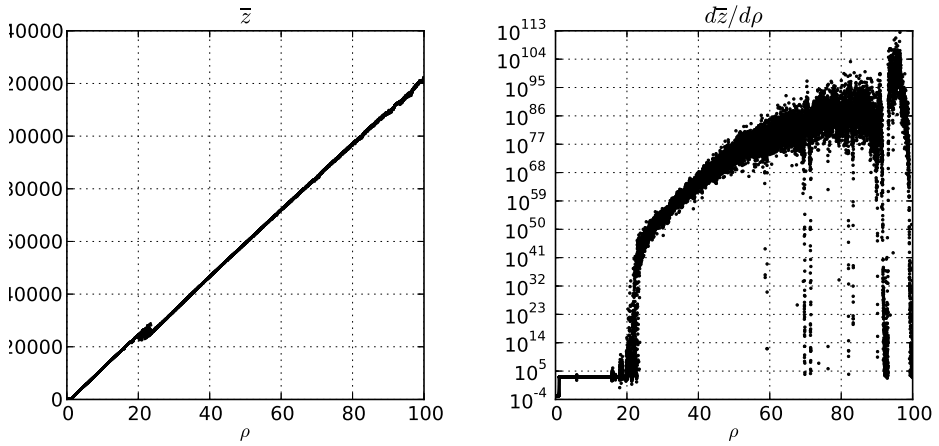
at a parameter s, where f is sufficiently smooth with respect to u and s. Consider an infinitesimal perturbation $s_\epsilon = s + \epsilon$. The solution u at the unperturbed parameter value s is no longer a solution at the perturbed parameter value s_ϵ, but rather an approximate solution. By the shadowing lemma [22], one can find a $u_\epsilon(t), t \in [0,\infty)$ that satisfies

$$\frac{du_\epsilon}{dt} = f(u_\epsilon, s_\epsilon)\ , \tag{4.2}$$



(a) $\overline{z}$ for $T = 2.26$

(b) Sensitivity of $\overline{z}$ for $T = 2.26$



(c) $\overline{z}$ for $T = 131.4$

(d) Sensitivity of $\overline{z}^{(T)}$ for $T = 131.4$

Fig. 3.1: Plots created with similar procedure to Lea et al[15] (permission granted). Left column: finite time average $\overline{z}$ plotted against a system parameter r; the infinite time average $\overline{z}$ has a slope of about 1. Right column: the sensitivity of the finite time average with respect to r, computed with method outlined in Section 2. As the length of time averaging T increases, these sensitivities do not approximate the sensitivity of the infinite time averaged $d\overline{z}/dr$, which is about 1. Note the order of magnitude of the y-axes.

such that u and u_ϵ are infinitesimally close under an infinitesimal time transformation $t \to t + \epsilon\sigma(t)$. Mathematically, both

$$\frac{\|u_\epsilon(t + \epsilon\sigma(t)) - u(t)\|}{\epsilon} \quad \text{and} \quad \left|\frac{d\sigma}{dt}\right|$$

are uniformly bounded for $t \in [0, \infty)$. Therefore, we can define the *tangent direction of the shadowing trajectory*

$$u'(t) = \frac{u_\epsilon(t + \epsilon\sigma(t)) - u(t)}{\epsilon} .$$

By taking its time derivative, substituting Equations (4.1) and (4.2), using chain rule, and ignoring $O(\epsilon^2)$ terms, we obtain the governing equation for u':

$$\frac{du'}{dt} = \frac{\partial f}{\partial u}u' + \frac{\partial f}{\partial s} + \frac{d\sigma}{dt}f , \quad 0 < t < T , \tag{4.3}$$

where $d\sigma/dt$ is the *rate of time dilation*.

By enabling uniform convergence and exchange of limit and derivative [19], the shadowing trajectory enables computing derivative of long time averaged quantities $\langle J \rangle$ as in Equation (3.1). Let the finite time average of the perturbed solution u_ϵ be

$$\overline{J}_\epsilon = \frac{1}{T + \epsilon\sigma(T)} \int_0^{T+\epsilon\sigma(T)} J(u_\epsilon(\tau), s) d\tau .$$

With change of variable $\tau = t + \epsilon\sigma(t)$, we obtain $d\tau = (1 + \epsilon\, d\sigma/dt)\, dt$, and

$$\overline{J}_\epsilon = \frac{1}{T + \epsilon\sigma(T)} \int_0^{T} J(u_\epsilon(t + \epsilon\sigma(t)), s) \left(1 + \epsilon\frac{d\sigma}{dt}\right) dt .$$

The difference between $\overline{J}_\epsilon$ and the unperturbed average $\overline{J} = \frac{1}{T}\int_0^T J(u, s)\, dt$ can be obtained via subtraction,

$$\frac{\overline{J}_\epsilon - \overline{J}}{\epsilon} = \overline{\frac{\partial J}{\partial u}u'} + \overline{\frac{\partial J}{\partial s}} + \overline{\frac{d\sigma}{dt}J} - \overline{\frac{d\sigma}{dt}}\,\overline{J} + O(\epsilon) , \tag{4.4}$$

where the over-line denotes time averaging in $[0, T]$.

It is worth noting that a uniformly shadowing trajectory u' exists for the whole $t \in [0, \infty)$ only when the attractor of the dynamical system is uniformly hyperbolic [23] [24]. A semi-infinite shadowing trajectory may not exist for quasi-hyperbolic and non-hyperbolic systems. However, the least squares procedure described in the next subsection often computes the tangent direction of a shadowing trajectory for sufficiently long time to permit accurate computation of sensitivities.

4.2. Least squares for computing the tangent direction of a shadowing trajectory. The tangent direction of the shadowing trajectory u' we solve the constrained linear least squares system

$$\min_{\eta(t), v(t)} \int_0^T \frac{1}{2}v^T v + \frac{\alpha^2}{2}\eta^2\, dt , \quad \text{s.t.} \quad \frac{dv}{dt} = \frac{\partial f}{\partial u}v + \frac{df}{ds} + \eta f . \tag{4.5}$$

The solution $v(t)$ and $\eta(t)$ in this problem approximate the tangent direction of the shadowing trajectory $u'(t)$ and the rate of time dilation $d\sigma/dt$, respectively. v approximates u' because their difference satisfies the equation

$$\frac{d}{dt}(v - u') = \frac{\partial f}{\partial u}(v - u') + \left(\eta - \frac{d\sigma}{dt}\right) f . \tag{4.6}$$

By decomposing $v - u'$ into Lyapunov eigenvectors $v - u' = \sum a_k(t)\phi_k(t)$, where each Lyapunov eigenvector satisfies

$$\frac{d\phi_k}{dt} = \frac{\partial f}{\partial u}\phi_k - \lambda_k \phi_k \ ,$$

where λ_k is the corresponding Lyapunov exponent. In particular, the Lyapunov exponent corresponding to the zero Lyapunov exponent $\lambda_{k_0} = 0$ is $\phi_{k_0} = f(u)$. By substituting the decomposition of $v - u'$ into the We find that the coefficients a_k of the decomposition satisfy the equation (4.6), we obtain

$$\frac{da_k}{dt} = \begin{cases} \lambda_k a_k & \lambda_k \neq 0 \\ \eta - d\sigma/dt & \lambda_k = 0 \end{cases}$$

For negative λ_k, a_k decays exponentially, thus $a_k(t) \ll a_k(0)$ for $t \gg 1/|\lambda_k|$. For positive λ_k, a_k decays exponentially as t goes backward from T, thus $a_k(t) \ll a_k(T)$ for $T - t \gg 1/|\lambda_k|$. Therefore, $v - u'$ is small for $t \in [T_B, T - T_E]$ other than a time transformation corresponding to a shift in the direction of $\phi_{k_0} = f(u)$. The buffer times are

$$T_B = -\frac{1}{\max\{\lambda_k | \lambda_k < 0\}} \ , \quad T_E = \frac{1}{\min\{\lambda_k | \lambda_k > 0\}} \ .$$

When $T \gg T_B + T_E$, the desired sensitivity derivative can be approximated by substituting v for u' and η for $d\sigma/dt$ in Equation (4.4), i.e.,

$$\frac{d\langle J \rangle(s)}{ds} \approx \overline{\frac{\partial J}{\partial u}} v + \overline{\frac{\partial J}{\partial s}} + \overline{\eta J} - \overline{\eta}\,\overline{J} \ . \tag{4.7}$$

4.3. Numerical solution of the least squares problem. To solve the constrained least squares system, we first convert it from a continuous-time problem to a discrete-time problem. Here we use the trapezoidal time integration rule to discretize d/dt. For a given parameter value s, suppose the solution of the nonlinear equation is obtained at $u_n = u(n\Delta t), n = 0, \ldots, N = T/\Delta t$. Denote the discrete solution as $v_n = v(n\Delta t)$ and $\eta_n = \eta((n - \frac{1}{2})\Delta t)$. The continuous least squares problem (4.5) can be discretized as

$$\begin{aligned} &\min_{v_n, \eta_n} \sum_{n=0}^{N} \frac{\|v_n\|_2^2}{2} + \alpha^2 \sum_{n=1}^{N} \frac{\eta_n^2}{2} \ , \qquad \text{s.t.} \\ &\frac{v_n - v_{n-1}}{\Delta t} = \frac{A_n v_n + A_{n-1} v_{n-1}}{2} + b_n + \eta_n f_n \ , \quad 1 \leq n \leq N \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} A_n &= \frac{\partial f}{\partial u}(u_n, s) \ , \\ b_n &= \frac{1}{2}\left(\frac{df}{ds}(u_n, s) + \frac{df}{ds}(u_{n-1}, s) \right) \ , \\ f_n &= \frac{f(u_n, s) + f(u_{n-1}, s)}{2} \end{aligned} \tag{4.9}$$

The optimality condition of this linear constrained least-squares problem can be represented by the following KKT system

$$\left[\begin{array}{ccccccccc|ccc} I & & & & & & & & & F_0^T & & \\ & \alpha^2 & & & & & & & & f_1^T & & \\ & & I & & & & & & & G_1^T & F_1^T & \\ & & & \alpha^2 & & & & & & & f_2^T & \\ & & & & I & & & & & & G_2^T & \\ & & & & & \ddots & & & & & \ddots & F_{N-1}^T \\ & & & & & & \alpha^2 & & & & & f_N^T \\ & & & & & & & I & & & & G_N^T \\ \hline F_0 & f_1 & G_1 & & & & & & & & & \\ & & F_1 & f_2 & G_2 & & & & & & & \\ & & & & \ddots & \ddots & & & & & & \\ & & & & & & F_{N-1} & f_N & G_N & & & \end{array}\right] \left[\begin{array}{c} v_0 \\ \eta_1 \\ v_1 \\ \eta_2 \\ v_3 \\ \vdots \\ \eta_N \\ v_N \\ \hline w_1 \\ w_2 \\ \vdots \\ w_N \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \hline -b_1 \\ -b_2 \\ \vdots \\ -b_N \end{array}\right] \tag{4.10}$$

where the matrices F_n and G_n are

$$F_n = \frac{I}{\Delta t} + \frac{A_n}{2}\,, \quad G_n = -\frac{I}{\Delta t} + \frac{A_n}{2}\,, \quad n = 0, \ldots, N\,. \tag{4.11}$$

This KKT system can be solved by Gauss elimination of the lower-left block, thereby forming the Schur complement of the KKT system

$$\mathbf{B}\mathbf{B}^T\mathbf{w} = \mathbf{b} \tag{4.12}$$

where

$$\mathbf{B} = \begin{bmatrix} F_0 & \frac{f_1}{\alpha} & G_1 & & & & \\ & F_1 & \frac{f_2}{\alpha} & G_2 & & & \\ & & & \ddots & \ddots & & \\ & & & & F_{N-1} & \frac{f_N}{\alpha} & G_N \end{bmatrix}\,, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}\,, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \tag{4.13}$$

This symmetric positive definite Schur complement matrix $\mathbf{B}\mathbf{B}^T$ is block-tri-diagonal, where the block size is the dimension of the dynamical system m. It can be solved with a banded direct solver, which takes $O(N\,m^3)$ floating point operations [25]. One can also perform a sparse QR factorization on the block-bi-diagonal $\mathbf{B}^T$, and then use backward and forward substitution to compute $\mathbf{w}$. The sparse QR factorization also takes $O(N\,m^3)$ floating point operations [25]. In addition, iterative solution methods can potentially be used to efficiently solve the system even for large m.

Once $\mathbf{m}$ is computed, v_n and η_n can be computed from the upper blocks of the KKT system (4.10). By setting $w_0 = w_{N+1} = 0$, we obtain

$$v_n = F_n^T w_{n+1} + G_n^T w_n\,, \quad n = 0, \ldots, N \qquad \eta_n = -\frac{f_n^T w_n}{\alpha^2}\,, \quad n = 1, \ldots, N \tag{4.14}$$

Once v_n and η_n are computed, the desired sensitivity derivative can then be computed by discretizing Equation (4.7)

$$\frac{d\langle J\rangle}{ds} \approx \sum_{n=0}^{N} \left(g_n^T v_n + \frac{1}{N+1}\frac{dJ}{ds}(u_n, s) \right) + \sum_{n=1}^{N} \eta_n h_n \tag{4.15}$$

where

$$
\begin{aligned}
g_n &= \frac{1}{N+1}\frac{\partial J}{\partial u}(u_n, s)\ , \quad n = 0, \ldots, N\ , \\
h_n &= \frac{1}{N}\left(\frac{J(u_n, s) + J(u_{n-1}, s)}{2} - \frac{1}{N+1}\sum_{i=0}^{N} J(u_i, s)\right), \quad n = 1, \ldots, N
\end{aligned}
\tag{4.16}
$$

4.4. Summary of the algorithm.

1. Choose a sufficiently small time step size Δt and sufficiently large time domain size $N\Delta t$.
2. Compute a solution to the equation (4.1) at $u_n = u(n\Delta t), n = 0, \ldots, N$.
3. Compute the vectors and matrices f_n, b_n, F_n and G_n as defined in Equations (4.9) and (4.11).
4. Choose a nonzero α, form matrix $\mathbf{B}$ and solve system (4.12) for $\mathbf{w}$.
5. Compute v_n and η_n from Equation (4.14).
6. Compute desired sensitivity using Equation (4.15).

The computational cost of algorithm is $O(N\, m^3)$, where N is the number of time steps and m is the dimension of the dynamical system.

4.5. Adjoint formulation of the sensitivity computation method. An alternative method of computing the same sensitivity estimate as in Equation (4.15) is through the corresponding discrete adjoint formulation. In this formulation, one first solves the adjoint system

$$
\left[\begin{array}{cccccccc|ccc}
I & & & & & & & & F_0^T & & \\
 & \alpha^2 & & & & & & & f_1^T & & \\
 & & I & & & & & & G_1^T & F_1^T & \\
 & & & \alpha^2 & & & & & & f_2^T & \\
 & & & & I & & & & & G_2^T & \\
 & & & & & \ddots & & & & \ddots & F_{N-1}^T \\
 & & & & & & \alpha^2 & & & & f_N^T \\
 & & & & & & & I & & & G_N^T \\
\hline
F_0 & f_1 & G_1 & & & & & & & & \\
 & & F_1 & f_2 & G_2 & & & & & & \\
 & & & & \ddots & \ddots & & & & & \\
 & & & & & F_{N-1} & f_N & G_N & & &
\end{array}\right]
\left[\begin{array}{c} \hat{v}_0 \\ \hat{\eta}_1 \\ \hat{v}_1 \\ \hat{\eta}_2 \\ \hat{v}_3 \\ \vdots \\ \hat{\eta}_N \\ \hat{v}_N \\ \hline \hat{w}_1 \\ \hat{w}_2 \\ \vdots \\ \hat{w}_N \end{array}\right]
=
\left[\begin{array}{c} g_0 \\ h_1 \\ g_1 \\ h_2 \\ g_2 \\ \vdots \\ h_N \\ g_N \\ \hline 0 \\ 0 \\ \vdots \\ 0 \end{array}\right]
\tag{4.17}
$$

where the sub-matrices F_n, G_n and vectors f_n, b_n are defined as in Equations (4.11), (4.9) and (4.16). This system (4.17) has exactly the same matrix as the system (4.10), but with a different right hand side. It can be solved by inverting

$$
\mathbf{B}\mathbf{B}^T\hat{\mathbf{w}} = \mathbf{B}\mathbf{g}\ , \tag{4.18}
$$

where $\mathbf{B}$ is defined in Equation (4.13), $\hat{\mathbf{w}} = (\hat{w}_1, \ldots, \hat{w}_N)$, and $\mathbf{g}$ is the upper portion of the right hand side of Equation (4.17). Once $\hat{\mathbf{w}}$ is computed, $d\overline{J}/ds$ can be computed via

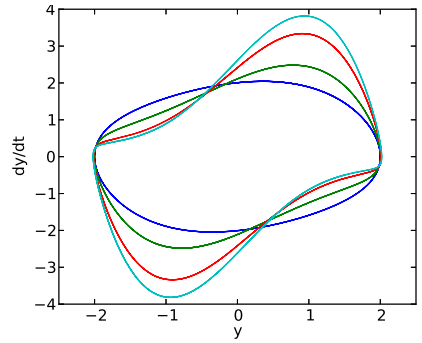
$$
\frac{d\overline{J}}{ds} \approx \sum_{n=1}^{N} b_n^T \hat{w}_n + \frac{1}{N+1}\sum_{n=0}^{N}\frac{dJ}{ds}(u_n, s)\ , \tag{4.19}
$$

where b_n is defined in Equation (4.9). It can be shown through matrix manipulation that the adjoint sensitivity estimate is equal to the sensitivity estimate in Section 4 up to round-off error. The examples in the following sections use the algorithm in Section 4.4.

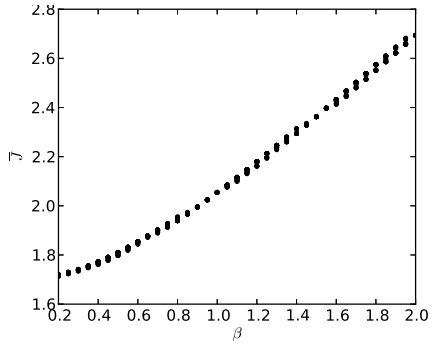
5. Application to the Van der Pol oscillator. We apply our method to the Van der Pol oscillator

$$\frac{d^2y}{dt^2} = -y + \beta(1-y^2)\frac{dy}{dt}\,. \tag{5.1}$$

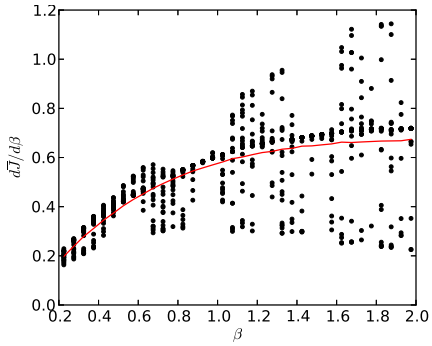
The parameter in the system is β. Figure 5.1a shows the limit cycle attractor as β



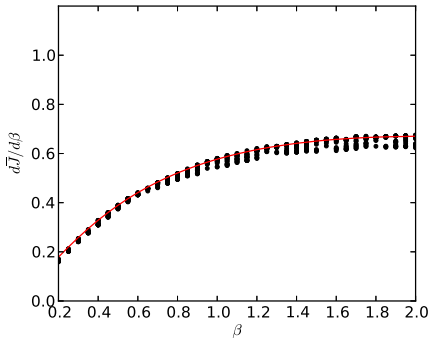
(a) Limit cycle attractor of the Van der Pol oscillator for $\beta = 0.2, 0.8, 1.6$ and 2.0.



(b) $\overline{J}$ as a function of β, computed on 20 trajectories of length 50 for each value of β.



(c) $d\overline{J}/d\beta$ as a function of β computed with finite difference with step size $\Delta\beta = 0.05$. The black dots are computed on 20 trajectories of length 50 at each value of β. The red line is computed on a trajectory of length 5000 for each β.



(d) $d\overline{J}/d\beta$ as a function of β computed with Least Squares sensitivity analysis. The black dots are computed on 20 trajectories of length 50 at each value of β. The red line is computed on a trajectory of length 5000 at each β.

Fig. 5.1: Least Squares Sensitivity Analysis of the van der Pol oscillator.

varies from 0.2 to 2.0. As β increases, the maximum magnitude of dy/dt significantly increases. Our objective function is chosen as the L^8 norm of dy/dt, which has a similar trend to the L^∞ norm and reflects the magnitude of the peak in dy/dt. Denoting $u = (u^{(1)}, u^{(2)}) = (y, dy/dt)$ as the state vector, we convert the second order ODE (5.1) into two coupled first order ODEs. The objective function can be written

as

$$\langle J \rangle = \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \left(\frac{dy}{dt} \right)^8 dt \right)^{\frac{1}{8}} = \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T u^{(2)\,8}\, dt \right)^{\frac{1}{8}} \tag{5.2}$$

The method described in Section 4 is then applied to compute v: for each β, we start the simulation at $t = -50$, with uniform $[0, 1]$ random initial conditions for both variables at $t = -50$. The ODE is integrated to $t = 0$ to ensure that $u(0)$ is almost on the limit cycle attractor. A trajectory $u(t), 0 \le t \le 50$ is then computed using a scipy[26] wrapper of lsoda[27], with time step size $\Delta t = 0.02$. The resulting $N + 1 = 2501$ states u_n along the trajectory are used to construct the linear system (4.10) for computing v_n.

The linearized solution v is then used to estimate the sensitivity of the output $\overline{J}$ with respect to the parameter β:

$$\frac{d\langle J \rangle}{d\beta} \approx \frac{\overline{J}^{-7}}{50} \int_0^{50} u^{(2)\,7} v^{(2)}\, dt \tag{5.3}$$

The computed sensitivity is compared against finite difference values in Figure 5.1. Both finite difference and least squares sensitivity analysis is repeated 20 times on randomly initiated short trajectories and a long trajectory for each value of β. The error in these calculations can be estimated from the spread of derivative values computed on the same parameter value. Figure 5.1 indicates that the least squares method is more accurate than finite difference for the same trajectory length.

6. Application to the Lorenz system. We apply our method to the Lorenz system (3.3), and analyze sensitivity to the parameter ρ in the system. The behavior of the Lorenz system as ρ changes from 0 to 100 is shown in Figure 6.1a, and can be summarized as following [28]:

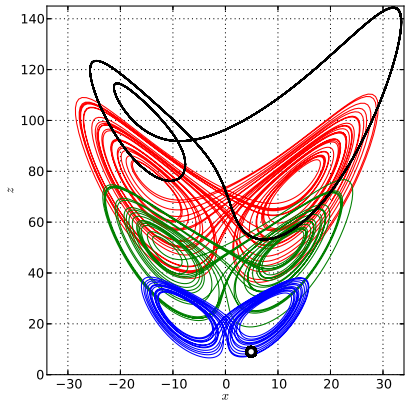
- Stable fixed point attractor at $(0, 0, 0)$ for $0 \le \rho <= 1$.
- Two stable fixed point attractors at $x = y = \pm\sqrt{\beta(\rho - 1)}, z = \rho - 1$ for $1 < \rho < 24.74$.
- Quasi-hyperbolic strange attractors for $24.06 < \rho < 31$. This includes the classic Lorenz attractor at $\rho = 28$.
- Non-hyperbolic quasi-attractors for $31 < \rho < 99.5$.
- Periodic limit cycle attractors with an infinite series of period doubling for $\rho > 99.5$.

Despite the many transitions in the fundamental nature of the system, the mean z value

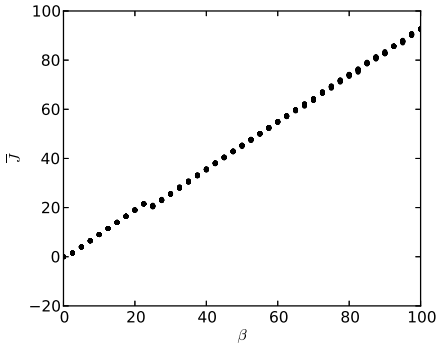
$$\langle z \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T z\, dt \tag{6.1}$$

apparently increases as the parameter ρ increases. $\langle z \rangle$ is chosen to be our time averaged output quantity in this study.

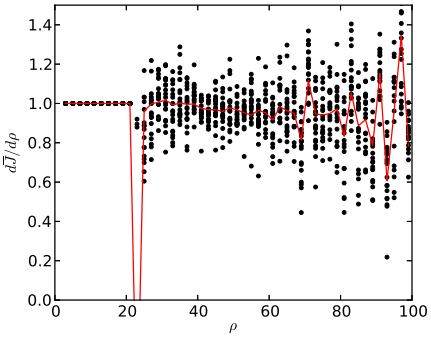
Denoting $u = (x, y, z)$, the method described in Section 4.4 is applied to compute v: for each ρ, we start the simulation at $t = -50$, with random uniform $[0, 1]$ as initial conditions for x, y and z. The Lorenz equation is integrated to $t = 0$ to ensure that $u(0)$ is almost on the attractor. A trajectory $u(t), 0 \le t \le 50$ is then computed using a scipy[26] wrapper of lsoda[27], with time step size $\Delta t = 0.01$. The resulting 5001 states along the trajectory are used to construct the linear system (4.10) for



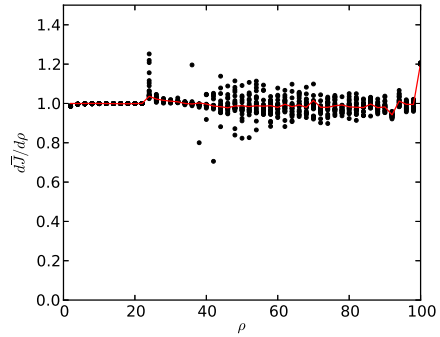
(a) Attractors of the Lorenz system for $\rho = 10$ (open circle), $\rho = 25, 50, 75$ and 100 (blue, green, red and black lines, respectively)



(b) $\overline{z}$ as a function of r, computed on 20 trajectories of length 50 for each value of ρ.



(c) $d\overline{J}/d\rho$ as a function of ρ computed with finite difference with step size $\Delta\rho = 2$. The black dots are computed on 20 trajectories of length 50 at each value of ρ. The red line is computed on a trajectory of length 5000 at each ρ.



(d) $d\overline{J}/d\rho$ as a function of ρ computed with Least Squares Sensitivity Analysis. The black dots are computed on 20 trajectories of length 50 at each value of ρ. The red line is computed on a trajectory of length 5000 at each ρ.

Fig. 6.1: Least Squares Sensitivity Analysis of the Lorenz system.

computing v. The computed linearized state v is plotted with the primal state u in Figure 6.1 for four different values of ρ. The different cases correspond to a fixed point attractor ($\rho = 10$), a quasi-hyperbolic strange attractor ($\rho = 25$), a non-hyperbolic attractor ($\rho = 75$) and a limit cycle attractor ($\rho = 100$).

The linearized solution v is then used to estimate the sensitivity of the output $\langle J \rangle := \langle z \rangle$ with respect to the parameter ρ:

$$\frac{d\langle J \rangle}{d\rho} \approx \frac{1}{50} \int_0^{50} (0, 0, 1)^T \, v \, dt \tag{6.2}$$

The computed sensitivity is compared against finite difference values in Figure 6.1. The dip in the finite difference value at around $\rho = 22.5$ is due to the discontinuous

transition from fixed point attractors to strange attractors at $24.0 \leq \rho \leq 24.74$ (note that within this small range, the two types of attractors co-exist). For $24.74 < \rho < 31$, the Lorenz system is dominated by a quasi-hyperbolic attractor. Least squares sensitivity analysis computes accurate and consistent gradients on randomly chosen short trajectories on the attractor.

As ρ increases beyond 31, the system is non-hyperbolic and its trajectories form an object known as a quasi-attractor [29]. For $\rho > 99.5$, the system transitions to periodic oscillations, then goes through an infinite series of period doubling bifurcations. The shadowing lemma on which the least squares sensitivity analysis is based no longer holds in these situations. However, we find that the method can still compute derivatives more accurate than finite difference method on the same trajectory lengths.

7. Application to an aero-elastic limit cycle oscillator. We apply our method to a simple model of aeroelastic limit cycle oscillation, as shown in Figure 7.1. The model is described in detail by Zhao and Yang[30]. The governing equations are

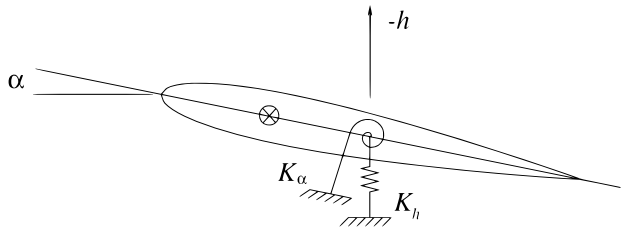


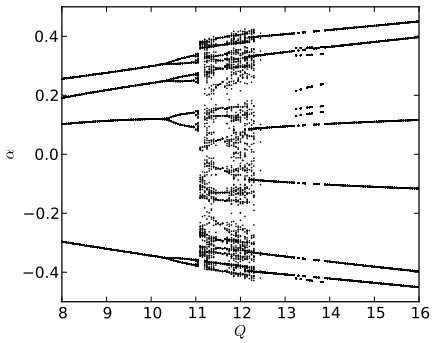
Fig. 7.1: Model aero-elastic oscillator

$$\begin{aligned} &\frac{d^2h}{dt^2} + 0.25\,\frac{d\alpha}{dt} + 0.1\,\frac{dh}{dt} + 0.2\,h + 0.1\,Q\,\alpha = 0 \\ &0.25\,\frac{d^2h}{dt^2} + 0.5\,\frac{d^2\alpha}{dt^2} + 0.1\,\frac{d\alpha}{dt} + 0.5\,\alpha + 20\,\alpha^3 - 0.1\,Q\,\alpha = 0 \end{aligned} \tag{7.1}$$

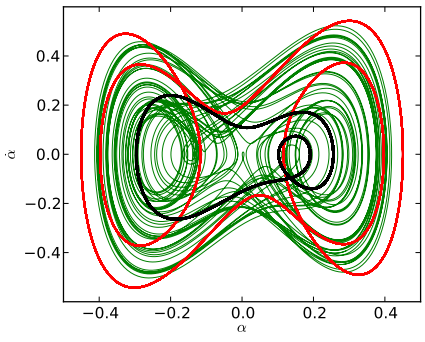
where h is the plunging degree of freedom, and α is the pitching degree of freedom. We analyze sensitivity to the reduced dynamic pressure Q.

The bifurcation diagram of α as Q increases from 8 to 16 is shown in Figure 7.2a. The behavior of the system as Q varies is complex [31]: At low values of Q, the system exhibits asymmetric periodic limit cycle oscillations. As Q increases beyond about 10.25, a series of period doubling bifurcations occurs, leading to transition into chaos just beyond $Q = 11$. At about $Q = 12.5$, the system transitions to symmetric periodic limit cycle oscillation. At above about $Q = 13.25$, there appears to be small windows in which the limit cycle oscillation transitions discontinuously back to asymmetric periodic shape. Finally, at about $Q = 13.9$, the system recovers symmetric periodic limit cycle oscillations. The phase plot of the system at several values of Q is shown in Figure 7.2b. These include an asymmetric periodic limit cycle oscillation at $Q = 8$, a chaotic limit cycle oscillation at $Q = 12$, and a symmetric periodic limit cycle oscillation at $Q = 16$.

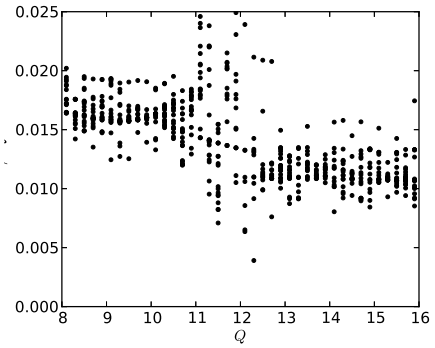
We observe that the magnitude of the oscillation grows as Q increases. Here we choose the L^8 norm of the pitch angle α as the objective function. The L^8 norm has similar trend as the L^∞ norm, which indicates the magnitudes of the oscillations in the



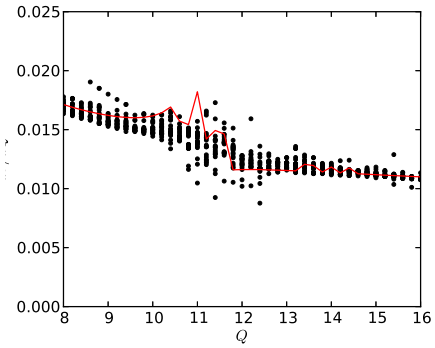
(a) Bifurcation diagram in the parameter range considered.



(b) Phase plots (α vs $\dot{\alpha} = d\alpha/dt$) at $Q = 8$ (black), $Q = 12$ (green) and $Q = 16$ (red).



(c) $d\overline{J}/dQ$ as a function of Q computed with finite difference with step size $\Delta Q = 0.2$. The black dots are computed on 20 trajectories of length 300 at each value of Q.



(d) $d\overline{J}/dQ$ as a function of Q computed with Least Squares Sensitivity Analysis. The black dots are computed on 20 trajectories of length 300 at each value of Q. The red line is computed on a trajectory of length 30000 at each Q.

Fig. 7.2: Least Squares Sensitivity Analysis on the aero-elastic oscillator model (7.1).

pitching degree of freedom. Denoting $u = (u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)}) = (y, \alpha, dy/dt, d\alpha/dt)$ as the state vector, we convert the pair of second order ODEs (7.1) into a set of four first order ODEs. The objective function can be written as

$$\langle J \rangle = \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \alpha^8 \, dt \right)^{\frac{1}{8}} = \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T u^{(2)\,8} \, dt \right)^{\frac{1}{8}} \tag{7.2}$$

We use the method described in Section 4.4 to compute the sensitivity derivative of the objective function to the input parameter Q. For each Q, we start the simulation at $t = -300$, with uniform $[0, 1]$ random initial conditions for all 4 variables at $t = -300$. The ODE is integrated to $t = 0$ to ensure that $u(0)$ is almost on the attractor. A trajectory $u(t), 0 \le t \le 300$ is then computed using a scipy[26] wrapper of lsoda[27], with time step size $\Delta t = 0.02$. The resulting 15001 states along the trajectory are used to construct the linear system (4.10) for computing v.

The linearized solution v is then used to estimate the sensitivity of the output $\overline{J}$

with respect to the parameter Q:

$$\frac{d\langle J\rangle}{dQ} \approx \frac{\overline{J}^{-7}}{300}\int_0^{300} u^{(2)\,7} v^{(2)}\,dt \tag{7.3}$$

The computed sensitivity is compared against finite difference values in Figure 7.2. The sensitivity derivative computed using least squares sensitivity analysis is more accurate than finite difference when the system exhibits periodic and chaotic limit cycle oscillations.

8. Conclusion. We presented a computational method for sensitivity analysis of long time averaged quantities in periodic and chaotic dynamical systems. The new method overcomes the difficulties that cause fundamental failure for traditional sensitivity analysis methods. The method is demonstrated on the periodic van der Pol oscillator, the chaotic Lorenz attractor, and a simple aero-elastic oscillation model that exhibits mixed periodic and chaotic behavior. These applications demonstrate the effectiveness of our new sensitivity computation algorithm in many complex nonlinear dynamics regimes. These include fixed points, limit cycles, quasi-hyperbolic and non-hyperbolic strange attractors.

Acknowledgements. The first author acknowledges AFOSR Award F11B-T06-0007 under Dr. Fariba Fahroo and a subcontract of the DOE PSAAP Program at Stanford.

REFERENCES

[1] Jameson, A., "Aerodynamic Design via Control Theory," *Journal of Scientific Computing*, Vol. 3, 1988, pp. 233–260.

[2] Reuther, J., Jameson, A., Alonso, J., Rimlinger, M., and Saunders, D., "Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers," *Journal of aircraft*, Vol. 36, No. 1, 1999, pp. 51–60.

[3] Bewley, T., "Flow control: new challenges for a new Renaissance," *Progress in Aerospace Sciences*, Vol. 37, No. 1, 2001, pp. 21–58.

[4] Bewley, T., Moin, P., and Temam, R., "DNS-based predictive control of turbulence: an optimal target for feedback algorithms," *J. Fluid Mech.*, Vol. 447, 2001, pp. 179–225.

[5] Allen, M. and Maute, K., "Reliability-based design optimization of aeroelastic structures," *Structural and Multidisciplinary Optimization*, Vol. 27, 2004, pp. 228–242.

[6] Stanford, B. and Beran, P., "Cost reduction techniques for the design of non-linear flapping wing structures," *International Journal for Numerical Methods in Engineering*, Vol. 88, No. 6, 2011, pp. 533–555.

[7] Stanford, B., Beran, P., and Kurdi, M., "Adjoint sensitivities of time-periodic nonlinear structural dynamics via model reduction," *Comput. Struct.*, Vol. 88, No. 19-20, Oct. 2010, pp. 1110–1123.

[8] Stanford, B. and Beran, P., "Computational strategies for reliability-based structural optimization of aeroelastic limit cycle oscillations," *Struct. Multidiscip. Optim.*, Vol. 45, No. 1, Jan. 2012, pp. 83–99.

[9] Wang, Q., Gleich, D., Saberi, A., Etemadi, N., and Moin, P., "A Monte Carlo method for solving unsteady adjoint equations," *Journal of Computational Physics*, Vol. 227, No. 12, 2008, pp. 61846205.

[10] Wang, Q., Moin, P., and Iaccarino, G., "Minimal Repetition Dynamic Checkpointing Algorithm for Unsteady Adjoint Calculation," *SIAM Journal on Scientific Computing*, Vol. 31, No. 4, 2009, pp. 2549–2567.

[11] Krakos, J., Wang, Q., Hall, S., and Darmofal, D., "Sensitivity analysis of limit cycle oscillations," *Journal of Computational Physics*, Vol. 231, No. 8, 2012, pp. 3228–3245.

[12] Yamaleev, N. K., Diskin, B., and Nielsen, E. J., "Local-in-time adjoint-based method for design optimization of unsteady flows," *Journal of Computational Physics*, Vol. 229, No. 14, 2010, pp. 5394 – 5407.

[13] Mani, K. and Mavriplis, D. J., "Unsteady Discrete Adjoint Formulation for Two-Dimensional Flow Problems with Deforming Meshes," *AIAA Journal*, Vol. 46, No. 6, 2008, pp. 1351–1364.
[14] Fidkowski, K. and Roe, P., "An Entropy Adjoint Approach to Mesh Refinement," *SIAM Journal on Scientific Computing*, Vol. 32, No. 3, 2010, pp. 1261–1287.
[15] Lea, D., Allen, M., and Haine, T., "Sensitivity analysis of the climate of a chaotic system," *Tellus*, Vol. 52A, 2000, pp. 523–532.
[16] Eyink, G., Haine, T., and Lea, D., "Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights," *Nonlinearity*, Vol. 17, 2004, pp. 1867–1889.
[17] Thuburn, J., "Climate sensitivities via a FokkerPlanck adjoint approach," *Quarterly Journal of the Royal Meteorological Society*, Vol. 131, No. 605, 2005, pp. 73–92.
[18] Abramov, R. V. and Majda, A. J., "Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems," *Nonlinearity*, Vol. 20, No. 12, 2007, pp. 2793.
[19] Wang, Q., "Forward and Adjoint Sensitivity Computation for Chaotic Dynamical Systems," *Journal of Computational Physics*, Vol. 235, No. 15, 2013, pp. 1–15.
[20] Bryson, A. and Ho, Y., *Applied Optimal Control: Optimization, Estimation, and Control*, John Wiley & Sons Inc, Hoboken, NJ, 1979.
[21] Strogatz, S. H., *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (Studies in Nonlinearity)*, Studies in nonlinearity, Westview Press, 1st ed., Jan. 2001.
[22] Pilyugin, S. and y Pilyugin, S., *Shadowing in dynamical systems*, Vol. 1706, Springer, 1999.
[23] Ruelle, D., "Differentiation of SRB states for hyperbolic flows," *Ergodic Theory and Dynamical Systems*, Vol. 28, No. 02, 2008, pp. 613–631.
[24] Dolgopyat, D., "On differentiability of SRB states for partially hyperbolic systems," *Inventiones Mathematicae*, Vol. 155, 2004, pp. 389–449, 10.1007/s00222-003-0324-5.
[25] Golub, G. H. and Loon, C. F. V., *Matrix Computations*, The Johns Hopkins Univ. Press, Baltimore, 1996.
[26] Jones, E., Oliphant, T., Peterson, P., et al., "SciPy: Open source scientific tools for Python," 2001–2013, http://www.scipy.org/.
[27] Petzold, L., "Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations," *SIAM Journal on Scientific and Statistical Computing*, Vol. 4, No. 1, 1983, pp. 136–148.
[28] Sparrow, C., *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*, Springer-Verlag, New York, 1982.
[29] Bonatti, C., Díaz, L., and Viana, M., *Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective*, Encyclopaedia of Mathematical Sciences, Springer, 2010.
[30] Zhao, L. and Yang, Z., "Chaotic motions of an airfoil with non-linear stiffness in incompressible flow," *Journal of Sound and Vibration*, Vol. 138, No. 2, 1990, pp. 245–254.
[31] Lee, B., Price, S., and Wong, Y., "Nonlinear aeroelastic analysis of airfoils: bifurcation and chaos," *Progress in Aerospace Sciences*, Vol. 35, No. 3, 1999, pp. 205–334.

Appendix A. Limit of derivative versus derivative of limit.

For a given function $f(x;s)$ of x parameterized by s, the limit of its derivative

$$\lim_{s\to\infty} \frac{df}{dx} \tag{A.1}$$

is not always equal to the derivative of the limit

$$\frac{d}{dx}\left(\lim_{s\to\infty} f(x,s)\right) \tag{A.2}$$

An intuitive counter example relevant to this paper is the following parameterized function

$$f(x,s) = \frac{1}{s}\sin e^{\lambda s}x \tag{A.3}$$

Because the derivative of the function

$$\frac{df}{dx}(x,s) = \frac{e^{\lambda s}}{s}\cos e^{\lambda s}x \tag{A.4}$$

goes to infinity as $s \to \infty$, the limit of its derivative

$$\lim_{s\to\infty} \frac{df}{dx} \tag{A.5}$$

does not exist. However, the limit of the function

$$\lim_{s\to\infty} f(x,s) \equiv 0 \; ; \tag{A.6}$$

therefore, the derivative of the limit

$$\frac{d}{dx}\left(\lim_{s\to\infty} f(x,s)\right) \equiv 0 \; . \tag{A.7}$$

Thus the limit of derivative is not equal to the derivative of limit.

Appendix B. Least squares solution of a linear differential equation. The least squares solution of a linear differential equation

$$\min_{u(t),-\infty<t<\infty} \|u\|_2 \; , \quad \text{s.t.} \quad \frac{du}{dt} = A\,u + b(t) \; , \quad 0 < t < T \; , \tag{B.1}$$

automatically forces the divergent and convergent components of the solution propagate in different time directions. To illustrate this property, consider the case when u is a scalar. In this case, the matrix A is also a scalar $A = \lambda$. Because Equation (B.1) is a quadratic minimization problem with linear constraints, the principle of superposition applies, and it is sufficient to consider the right hand side $b(t) = \delta(t)$ being a Dirac delta function. The least squares solution satisfies

$$\min_{u(t),-\infty<t<\infty} \|u\|_2 \; , \quad \text{s.t.} \quad \frac{du}{dt} = \lambda\,u + \delta(t) \; , \quad 0 < t < T \; , \tag{B.2}$$

One can verify that for $\lambda > 0$, the solution to Equation (B.2) is

$$u(t) = \begin{cases} -e^{\lambda t} & t < 0 \\ 0 & t \geq 0 \end{cases} \tag{B.3}$$

while for $\lambda < 0$, the solution to Equation (B.2) is

$$u(t) = \begin{cases} e^{\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases} \tag{B.4}$$

These solutions indicates that the impulse response of an "unstable" equation ($\lambda > 0$) propagates backwards in time; while the impulse response of a "stable" equation ($\lambda < 0$) propagates forward in time.

For vector equations containing both stable and unstable modes, the least squares solution automatically separates each mode, allowing them to propagate in different time directions. We consider two examples. For simplicity consider $\lambda > 0$.

$$\min_{u(t),-\infty<t<\infty} \|u\|_2 \; , \quad \text{s.t.} \quad \frac{du}{dt} = \begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix} u + \begin{bmatrix} \delta(t) \\ \delta(t) \end{bmatrix} \; , \quad 0 < t < T \; , \tag{B.5}$$

$$\min_{u(t),-\infty<t<\infty} \|u\|_2 \; , \quad \text{s.t.} \quad \frac{du}{dt} = \begin{bmatrix} 0 & \lambda \\ \lambda & 0 \end{bmatrix} u + \begin{bmatrix} 2\delta(t) \\ 0 \end{bmatrix} \; , \quad 0 < t < T \; , \tag{B.6}$$

In the first example, it is easy to verify that the solution is a superposition of the solution of the stable and unstable scalar problem:

$$\begin{aligned} u_1(t) &= \begin{cases} e^{\lambda t} & t < 0 \\ 0 & t \geq 0 \end{cases} \\ u_2(t) &= \begin{cases} e^{-\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases} \end{aligned} \tag{B.7}$$

In the second example, the solution can be verified as

$$\begin{aligned} u_1(t) &= \begin{cases} e^{-\lambda t} & t > 0 \\ -e^{\lambda t} & t \geq 0 \end{cases} \\ u_2(t) &= \begin{cases} -e^{-\lambda t} & t > 0 \\ -e^{\lambda t} & t \leq 0 \end{cases} \end{aligned} \tag{B.8}$$

We can see that in both cases, the impulse response of the system is nonzero in both forward and backward time directions. Different modes corresponding to different rate of growth or decay propagates in different directions.